CAGRID PORTAL V. 2.0.2

Installation Guide

NATIONAL CANCER INSTITUTE®

Center for Bioinformatics

This is a U.S. Government work.

June 30, 2008

caBIG™ cancer Biomedical Informatics Grid™

Model caBIG™ Open Source Software License
v.2
Release Date: January 7, 2008

Copyright Notice. Copyright 2008 The Ohio State University Research Foundation (OSURF), Argonne National Labs (ANL), SemanticBits LLC (SemanticBits), and Ekagra Software Technologies Ltd. (Ekagra) ("caBIG™ Participant"). The caGrid 1.2 software was created with NCI funding and is part of the caBIG™ initiative. The software subject to this notice and license includes both human readable source code form and machine readable, binary, object code form (the "caBIG™ Software").

This caBIG™ Software License (the "License") is between caBIG™ Participant and You. "You (or "Your") shall mean a person or an entity, and all other entities that control, are controlled by, or are under common control with the entity. "Control" for purposes of this definition means (i) the direct or indirect power to cause the direction or management of such entity, whether by contract or otherwise, or (ii) ownership of fifty percent (50%) or more of the outstanding shares, or (iii) beneficial ownership of such entity.

License. Provided that You agree to the conditions described below, caBIG™ Participant grants You a non-exclusive, worldwide, perpetual, fully-paid-up, no-charge, irrevocable, transferable and royalty-free right and license in its rights in the caBIG™ Software, including any copyright or patent rights therein, to (i) use, install, disclose, access, operate, execute, reproduce, copy, modify, translate, market, publicly display, publicly perform, and prepare derivative works of the caBIG™ Software in any manner and for any purpose, and to have or permit others to do so; (ii) make, have made, use, practice, sell, and offer for sale, import, and/or otherwise dispose of caBIG™ Software (or portions thereof); (iii) distribute and have distributed to and by third parties the caBIG™ Software and any modifications and derivative works thereof; and (iv) sublicense the foregoing rights set out in (i), (ii) and (iii) to third parties, including the right to license such rights to further third parties. For sake of clarity, and not by way of limitation, caBIG™ Participant shall have no right of accounting or right of payment from You or Your sublicensees for the rights granted under this License. This License is granted at no charge to You. Your downloading, copying, modifying, displaying, distributing or use of caBIG™ Software constitutes acceptance of all of the terms and conditions of this Agreement. If you do not agree to such terms and conditions, you have no right to download, copy, modify, display, distribute or use the caBIG™ Software.

1. Your redistributions of the source code for the caBIG™ Software must retain the above copyright notice, this list of conditions and the disclaimer and limitation of liability of Article 6 below. Your redistributions in object code form must reproduce the above copyright notice, this list of conditions and the disclaimer of Article 6 in the documentation and/or other materials provided with the distribution, if any.
2. Your end-user documentation included with the redistribution, if any, must include the following acknowledgment: "This product includes software developed by the Ohio State University Research Foundation (OSURF), Argonne National Labs (ANL), SemanticBits LLC (SemanticBits), and Ekagra Software Technologies Ltd. (Ekagra)." If You do not include such end-user

documentation, You shall include this acknowledgment in the caBIG™ Software itself, wherever such third-party acknowledgments normally appear.

3. You may not use the names “The Ohio State University Research Foundation”, "OSURF", “Argonne National Labs”, “ANL”, “SemanticBits LLC”, "SemanticBits", “Ekagra Software Technologies Ltd.”, “Ekagra”, “The National Cancer Institute”, “NCI”, “Cancer Bioinformatics Grid” or “caBIG™” to endorse or promote products derived from this caBIG™ Software. This License does not authorize You to use any trademarks, service marks, trade names, logos or product names of either caBIG™ Participant, NCI or caBIG™, except as required to comply with the terms of this License.
4. For sake of clarity, and not by way of limitation, You may incorporate this caBIG™ Software into Your proprietary programs and into any third party proprietary programs. However, if You incorporate the caBIG™ Software into third party proprietary programs, You agree that You are solely responsible for obtaining any permission from such third parties required to incorporate the caBIG™ Software into such third party proprietary programs and for informing Your sublicensees, including without limitation Your end-users, of their obligation to secure any required permissions from such third parties before incorporating the caBIG™ Software into such third party proprietary software programs. In the event that You fail to obtain such permissions, You agree to indemnify caBIG™ Participant for any claims against caBIG™ Participant by such third parties, except to the extent prohibited by law, resulting from Your failure to obtain such permissions.
5. For sake of clarity, and not by way of limitation, You may add Your own copyright statement to Your modifications and to the derivative works, and You may provide additional or different license terms and conditions in Your sublicenses of modifications of the caBIG™ Software, or any derivative works of the caBIG™ Software as a whole, provided Your use, reproduction, and distribution of the Work otherwise complies with the conditions stated in this License.
6. THIS caBIG™ SOFTWARE IS PROVIDED "AS IS" AND ANY EXPRESSED OR IMPLIED WARRANTIES (INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE) ARE DISCLAIMED. IN NO EVENT SHALL THE OHIO STATE UNIVERSITY RESEARCH FOUNDATION ("OSURF"), ARGONNE NATIONAL LABS (“ANL”), SEMANTICBITS LLC ("SEMANTICBITS"), AND EKAGRA SOFTWARE TECHNOLOGIES LTD. (EKAGRA) OR ITS AFFILIATES BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS caBIG™ SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Revision History

The following is the revision history for this document.

Date	Version	Description	Revised By
June 28, 2008	2.0.2.1	Added caGrid license v2 and formatted.	Carolyn Kelley Klinger
June 27, 2008	2.0.2	Updated for caGrid Portal 2.0.2.	Manav Kher
December 18, 2007	2.0.1	Updated for caGrid Portal 2.0.1.	Eddie VanArsdall

Contents

About This Guide

This guide explains how to install, deploy, configure, and run the caGrid Portal (version 2.0.2) application. It also provides prerequisite requirements for the installation.

About the caGrid Portal

The caGrid Portal is a Web-based application that enables end users to discover and interact with the services that are available on the caGrid infrastructure. The portal serves as the primary visualization tool for the caGrid middleware and provides a standards-based platform for hosting caBIG-related tools. The portal also serves as a caBIG information source. Users of the portal have instant access to information about caBIG participants, caGrid points of contact (POCs), and caGrid-related news and events.

Audience for This Guide

This installation guide is intended for caGrid service developers, client application developers, and service administrators. Users of this guide should be familiar with Java and other programming languages, database concepts, and the Internet. This guide assumes that users intending to use caGrid resources in software applications have experience with building and using complex data systems.

Pre-Installation

Obtain the caGrid Portal Software

You can obtain the caGrid Portal software by downloading the release or checking out the project from CVS.

Downloading the release

To download the release, follow this link:

https://gforge.nci.nih.gov/frs/download.php/4201/cagrid-portal-2.0.2.zip

Checking out the project from CVS

To check out the project from CVS, use the settings in Table 1:

Setting	***Value***
Username	anonymous
Password	anonymous
Protocol	ssh
Host	cbiocvs2.nci.nih.gov
Repository	/share/content/gforge/cagrid-1-0
module	cagrid-1-0/Applications/cagrid-portal
tag	cagrid-portal-2-0-2_final

Table 1. Settings for checking out source code from CVS

If you are using a command line CVS client, use this command:

```
export CVS_RSH=ssh

cvs -d :ext:anonymous@cbiocvs2.nci.nih.gov:/share/content/gforge/cagrid-1-0 co \

       -r cagrid-portal-2-0-2_final \

       -d cagrid-portal cagrid-1-0/Applications/cagrid-portal
```

After checkout, the caGrid Portal source code directory (referred to henceforth as $SRC) will be located under `cagrid-portal` in the current directory.

Obtain the Required Software

Before installing the caGrid Portal, make sure that you have the following applications. Set the environment variables as specified.

Application/Version	Preparation
Java 1.5	Set `JAVA_HOME` and make sure that the Java SDK executable is on the `PATH`.
Ant 1.6.5	Set `ANT_HOME` and make sure that the Ant executable is on the `PATH`.
MySQL 5+	You will need privileges to create and delete databases.

Table 2. caGrid Portal installation requirements

Create the Databases

The caGrid Portal application requires two databases: one for Liferay data, and one for caGrid Portal data. By convention, the names of these databases are `lportal` and `portal2`. If you are using these database names, then you need to execute the following SQL in your MySQL database:

```
create database lportal character set utf8;

create database portal2;
```

You also need to provision an account that has access to these databases. If the same account will have full access to both databases, use the following SQL:

```
grant all privileges on lportal.* to 'portalacct'@'%' identified by 'mypwd';

grant all privileges on portal2.* to 'portalacct'@'%' identified by 'mypwd';

flush privileges;
```

This gives the user `portalacct` all privileges on these databases. This user can access the databases from any host. For more information, see the MySQL admin documentation at http://dev.mysql.com/doc/refman/5.0/en/index.html.

Create the SSL Certificate and Keystore

Some pages in the portal need to be protected with HTTPS. You need to create an SSL certificate that the embedded Tomcat instance (running in JBoss) will use. The installation script will configure the Tomcat HTTPS Connector, but you still must either create a certificate and PKCS12 keystore or specify the path to an existing keystore and provide the keystore password.

To create a keystore using the Java keytool, run this command:

```
$JAVA_HOME/bin/keytool -genkey -alias tomcat -keyalg RSA -keystore
/path/to/my/keystore
```

Make sure that you use the same password for the keystore and the key. When prompted for the first and last name, specify the host name.

For complete instructions on using keytool, follow this link: http://java.sun.com/j2se/1.5.0/docs/tooldocs/windows/keytool.html

For instructions on configuring JBoss and Tomcat to use SSL, follow these links:

- http://tomcat.apache.org/tomcat-5.5-doc/ssl-howto.html
- http://docs.jboss.org/jbossas/jboss4guide/r5/html/ch9.chapt.html#ch9.https.sect

By default, the installation script will assume that the keystore is located at `$HOME/portal-liferay/portal-keystore` and that the keystore password is `portal`. You can configure this location and password in the properties file that the installation script uses (described later).

Obtain a Google Maps API Key

You can get a Google Maps API key at http://www.google.com/apis/maps/signup.html. If your host name is `my.host.com` and the HTTP server is listening to port 8080, then the URL you should use is http://my.host.com:8080. Save this key for future use.

Obtain Yahoo ApplicationID

Get an ApplicationID from Yahoo at http://developer.yahoo.com/wsregapp . Save this for future use.

Configure the caGrid Trust Fabric

The caGrid Portal uses GAARDS to authenticate users, so you must configure the caGrid trust fabric on the machine that will host the portal. The portal itself will use the GTS client to maintain the trust fabric, but you must also bootstrap the trust fabric.

By default, the portal will use the `nci_prod` as the target grid. If you are using this grid, you don't need to do anything.

If you are using one of the following grids, the trust synchronization configuration has already been provided:

- `nci_dev`
- `nci_stage`
- `osu_dev`
- `training`

You will simply need to create a corresponding `build.properties` file. Look at `build-nci_qa.properties` as an example. When you run the installation script, you must specify the name of your target environment. For more information, see the next section, *Configure the caGrid Portal* Installation.

If you are using another target grid, then you need to do three things:

7. Create a `sync-description.xml` file to configure the GTS client that the portal uses.
8. Bootstrap the trust fabric by placing root certificates under the `$HOME/.globus/certificates` directory.
9. Configure the caGrid Portal to use your `sync-description.xml` configuration.

To configure the portal to uses your `sync-description.xml` file and certificates, you need to edit the `aggr.trust.syncgts.file` and `aggr.trust.certs.dir` properties to the path to your `sync-description.xml` file and the directory in which the root certificates are found, respectively.

To find directions for configuring a trust fabric using caGrid tools, follow this link: http://www.cagrid.org/mwiki/index.php?title=GTS:1.1:Administrators_Guide:Syncing_With_the_Trust_Fabric

Configure the caGrid Portal Installation

The caGrid Portal installation script is at `$SRC/build.xml`. This is an Ant build file configured by the properties defined in the `build.properties` file in the same directory. To customize the installation, you can directly edit `build.properties`, or you can override those properties by adding them to the `build-local.properties` file.

If you want to maintain installation configurations for multiple deployment tiers, you can create multiple properties files with the following naming convention: `build-<tier>.properties`, where `<tier>` is replaced with the name of the tier. For example, if you create a configuration for the `testing` tier, then you would create a file named `build-testing.properties`, and then run Ant from the `$SRC` directory as follows:

```
ant -target.env=testing install
```

See the `$SRC/build.properties` file itself for descriptions of all the properties. If you are using the `nci_prod target` grid and default installation location, then you will usually only need to edit the following properties:

- `liferay.admin.password`
- `liferay.db.host`
- `liferay.db.port`
- `liferay.db.name`
- `liferay.db.username`
- `liferay.db.password`
- `cagrid.portal.admin.email`
- `cagrid.portal.db.url`
- `cagrid.portal.db.username`
- `cagrid.portal.db.password`
- `cagrid.portal.geocoder.yahoo.appId`
- `cagrid.portal.map.google.apiKey`
- `cagrid.portal.security.encryption.key`

Installation

From the `$SRC` directory, run the following command:

```
  ant -Dtarget.env=<envname> install
```

If you have just directly updated `build.properties` or `build-local.properties`, then you would run the following command:

```
ant install
```

Post-Installation

Set Up the Environment

On Windows, use the System application in the Control Panel to set the `JBOSS_HOME` environment variable to point to the directory in which JBoss was installed. By default, the variable will point to the following path:

```
%HOMEDRIVE%%HOMEPATH%\portal-liferay\jboss-4.0.5.GA
```

On Unix/Linux/Mac, the default location will be `$HOME/portal-liferay/jboss-4.0.5.GA`. You can set the environment variable in the bash shell as follows:

```
export JBOSS_HOME=$HOME/portal/liferay/jboss-4.0.5.GA
```

Start the Application Server

On Windows, do the following:

1. Navigate to `%JBOSS_HOME%\bin`.
2. Double-click `run.bat`.

On Unix/Linux/Mac, do something like this:

```
cd $JBOSS_HOME/bin

chmod u+x *

./run.sh > portal.log &
```

Import the Site Structure

You can configure the basic site structure by using the Liferay administrative portlets to import Liferay Archive (`lar`) files. These files are available in the `$SRC/portals/liferay/lars/` directory.

To import the `lar` files, follow these steps:

Step	***Action***
1.	Using your browser, go to: https://<hostname>:8443.
2.	Sign in using the following parameters: • **username:** `portaladmin@cabig.nci.nih.gov` • **password:** `p0rtal@dmin`
3.	In the upper right corner, click the **Welcome** drop-down menu, then select **My Places** > **My Community**.
4.	Click the icon to the right of the **Private Pages** text.
5.	Click the **Import/Export** tab.
6.	Click the **Import** sub tab.
7.	Select the **Portlet Preferences** and **Portlet Data** checkboxes.
8.	Click the **Browse** button to navigate to and select `portaladmin-private-community.lar`.
9.	Click **OK**.
10.	Click **Import**.
11.	Click the back arrow icon in the upper right corner.
12.	In the upper right-hand corner, click the **Welcome** drop-down menu, then select **My Places** > **My Community** > **Private Pages**. **Note:** This time, just click the **Private Pages** text—not the icon to the right.
13.	Using the Communities portlet, click the **Communities I have joined** tab.
14.	Click the **Configure Pages** icon for the **Guest** community. This is the third icon from the left showing an image of two pieces of paper.

Step	***Action***
15.	Click the **Import/Export** tab.
16.	Click on the **Import** sub tab.
17.	Select the **Portlet Preferences** and **Portlet Data** check boxes.
18.	Click the **Browse** button to navigate to and select `guest-community.lar`.
19.	Click **OK**.
20.	Click **Import**.
21.	Click the back arrow icon in the upper, right corner of page.
22.	In the upper, right corner, click the **Welcome** drop-down menu, then select **My Places** > **Guest Community** > **Public Pages**. The caGrid Portal home page appears.
23.	Sign out by selecting the **Sign Out** option from the **Welcome** drop-down menu.

NOTE:

If you see an error message just below the Google Map on the Home page stating that you do not have privileges to view this portlet, simply restart JBoss.

Secure the Encryption Key

A file named `cagridportal.properties` will be generated and placed in two locations on the file system:

- `$JBOSS_HOME/server/default/deploy/liferay-portal.war/WEB-INF/classes/cagridportal.properties`
- `$JBOSS_HOME/server/default/deploy/cagridportlets/WEB-INF/classes/cagridportal.properties`

The value of the `cagrid.portal.security.encryption.key` property in this file will be used to encrypt authentication tickets as well as portal users' temporary grid credentials in the database. These files must be protected so that users' grid credentials cannot be decrypted by a malicious user who has access to both the hosting system and database. Set file permissions appropriately for your system.

Re-Installation

This section describes the steps needed to wipe out an existing installation and re-install. All existing data will be destroyed. See the Administrator's Guide for directions on backing up data and running batch imports of data.

1. Stop the JBoss application server.
2. Drop the `portal2` and `lportal` databases.
3. Re-create the databases.
4. From `$SRC`, run the following command:

```
    ant -Dtarget.env=<env> clean install
```

Firewall/Connectivity Considerations

By default, the portal installation script will download JBoss and several Liferay artifacts. If you are behind a firewall, you will need to provide the proxy configuration for your Java Virtual Machine. Essentially, you simply need to set the `ANT_OPTS` environment variable to include the standard Java proxy settings using the following command:

```
export ANT_OPTS="-Dhttp.proxyHost=proxy -Dhttp.proxyPort=8080"
```

For additional instructions, follow this link: http://ant.apache.org/manual/proxy.html.

IF YOU HAVE NO INTERNET ACCESS, you can still use the installation script. You will simply need to download the dependencies manually and then edit the following properties in `build.properties` so that they point to local directories and files:

- `liferay.jboss.home`
- `liferay.jboss.zip`
- `liferay.downloads.dir`
- `liferay.dependencies.zip`
- `liferay.war`

See `build-liferay.xml` for details.